Math 170S: Homework 4

Owen Jones

11/7/2023

Problem 1. Since each $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2), Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$, and $W_i \sim \mathcal{N}(\mu_W, \sigma_W^2)$, it follows that $\sum_{i=1}^{n_X} X_i \sim \mathcal{N}(n_X\mu_X, n_X\sigma_X^2)$, $\sum_{i=1}^{n_Y} Y_i \sim \mathcal{N}(n_Y\mu_Y, n_Y\sigma_Y^2)$, and $\sum_{i=1}^{n_W} W_i \sim \mathcal{N}(n_W\mu_W, n_W\sigma_W^2)$ by the linearity of mean and variance.
Using $E[nX] = nE[X]$ and $Var[nX] = nVar[\frac{X}{n}]$ $\overline{X} = \frac{1}{n_X}\sum_{i=1}^{n_X} X_i \sim \mathcal{N}(\mu_X, \frac{\sigma_X^2}{n_X})$, $\overline{Y} = \frac{1}{n_Y}\sum_{i=1}^{n_Y} Y_i \sim \mathcal{N}(\mu_Y, \frac{\sigma_Y^2}{n_Y})$, and $\overline{W} = \frac{1}{n_W}\sum_{i=1}^{n_W} W_i \sim \mathcal{N}(\mu_W, \frac{\sigma_W^2}{n_W})$.
By the linearity of mean and variance, $\overline{X} - \overline{Y} - \overline{W} \sim \mathcal{N}(\mu_X - \mu_Y - \mu_W, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y} + \frac{\sigma_W^2}{n_W})$. It follows $\frac{\overline{X}-\overline{Y}-\overline{W}-(\mu_X-\mu_Y-\mu_W)}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_W^2}{n_W}}} \sim \mathcal{N}(0,1)$. Choose $z_{\frac{\alpha}{2}}$ s.t $P(|\frac{\overline{X}-\overline{Y}-\overline{W}-(\mu_X-\mu_Y-\mu_W)}{\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_W^2}{n_W}}}| < z_{\frac{\alpha}{2}}) = 1-\alpha$. Thus, $((\overline{X}-\overline{Y}-\overline{W}) - z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_W^2}{n_W}}, (\overline{X}-\overline{Y}-\overline{W}) + z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_X^2}{n_X}+\frac{\sigma_Y^2}{n_Y}+\frac{\sigma_W^2}{n_W}})$ is a $100(1-\alpha)\%$ confidence interval.

Problem 2. We know $(\overline{X} - \overline{Y} - t_{\frac{\alpha}{2}}^{(4)}\sqrt{\frac{s_x^2+s_y^2}{5}}, \overline{X} - \overline{Y} + t_{\frac{\alpha}{2}}^{(4)}\sqrt{\frac{s_x^2+s_y^2}{5}})$ is a confidence interval for the difference of two means with unknown variance. $\overline{X} - \overline{Y} = \frac{1}{5}\sum_{i=1}^{n} x_i - y_i = -20.2$, $s_x^2 + s_y^2 = \frac{\sum_{i=1}^{5}(x_i - \overline{X})^2 + (y_i - \overline{Y})^2}{4} = 11986.5$, $\sqrt{\frac{s_x^2+s_y^2}{5}} = 48.96$. $t_{0.05}^{(4)} = 2.13185$. Thus, we obtain the confidence interval $(-124.57997, 84.17997)$.

Problem 3.

1. $\hat{p} = \frac{24}{642} \Rightarrow (\frac{24}{642} - 1.960 \cdot \sqrt{\frac{\frac{24}{642}(1-\frac{24}{642})}{642}}, \frac{24}{642} + 1.960 \cdot \sqrt{\frac{\frac{24}{642}(1-\frac{24}{642})}{642}}) \Rightarrow (0.02271, 0.05206)$ gives us an approximate 95% confidence interval for p.
2. $\hat{p} = \frac{24}{642} \Rightarrow (\frac{\frac{24}{642}+\frac{(1.960)^2}{2\cdot 642}-1.960\sqrt{\frac{\frac{24}{642}(1-\frac{24}{642})}{642}+\frac{(1.960)^2}{4*642^2}}}{1+\frac{(1.960)^2}{642}}, \frac{\frac{24}{642}+\frac{(1.960)^2}{2\cdot 642}+1.960\sqrt{\frac{\frac{24}{642}(1-\frac{24}{642})}{642}+\frac{(1.960)^2}{4*642^2}}}{1+\frac{(1.960)^2}{642}})$
$\Rightarrow (0.02525, 0.05502)$ is a 95% confidence interval for p.

Problem 4.

- We want to choose n s.t $z_{0.025}\sqrt{\frac{p(1-p)}{n}} \leq 0.03$. $z_{0.025}\sqrt{\frac{p(1-p)}{n}} \leq z_{0.025}\frac{0.5}{\sqrt{n}}$, so $\frac{z_{0.025}}{2\cdot 0.03} \leq \sqrt{n}$. Since both sides are positive, $(\frac{z_{0.025}}{2\cdot 0.03})^2 \leq n \Rightarrow$ choose $1068 = n$.
- We want to choose n s.t $z_{0.025}\sqrt{\frac{p(1-p)}{n}} \leq 0.02$. $z_{0.025}\sqrt{\frac{p(1-p)}{n}} \leq z_{0.025}\frac{0.5}{\sqrt{n}}$, so $\frac{z_{0.025}}{2\cdot 0.02} \leq \sqrt{n}$. Since both sides are positive, $(\frac{z_{0.025}}{2\cdot 0.02})^2 \leq n \Rightarrow$ choose $2401 = n$.

Problem 5. $\overline{X} - \overline{Y} \pm z_{0.05}\sqrt{\frac{\sigma_X^2+\sigma_Y^2}{n}}$ is a 90% confidence interval for $\mu_x - \mu_y$. We want to choose the smallest n s.t $z_{0.05}\sqrt{\frac{\sigma_X^2+\sigma_Y^2}{n}} \leq 4$. Because both sides are positive we want $\frac{z_{0.05}^2(15^2+25^2)}{16} \leq n \Rightarrow 144 \leq n$, so we choose $n = 144$